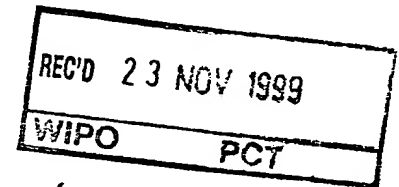


EP 99/7055



4

## Bescheinigung

Herr Hassan J o m a a in Gießen/Deutschland hat eine Patentanmeldung unter der Bezeichnung

"Gene des 1-Desoxy-D-xylulose-Biosynthesewegs"

am 21. Mai 1999 beim Deutschen Patent- und Markenamt eingereicht und erklärt, daß er dafür die Innere Priorität der Anmeldung in der Bundesrepublik Deutschland vom 22. September 1998, Aktenzeichen 198 43 279.8, in Anspruch nimmt.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig die Symbole C 07 K, C 12 N und C 12 Q der Internationalen Patentklassifikation erhalten.

München, den 2. November 1999

Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag

Sieck

Aktenzeichen: 199 23 567.8

**PRIORITY  
DOCUMENT**

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

Gene des 1-Desoxy-D-xylulose-Biosynthesewegs

Die vorliegende Erfindung betrifft DNA-Sequenzen, die bei Integration in das Genom von Viren, Eukaryonten und Prokaryonten die Isoprenoid-Biosynthese verändern sowie gentechnologische Verfahren zur Herstellung dieser transgenen Viren, Eukaryonten und Prokaryonten. Außerdem betrifft sie Verfahren zur Identifizierung von Stoffen mit herbizider, antimikrobieller, antiparasitärer, antiviraler, fungizider, bakterizider Wirkung bei Pflanzen und antimikrobieller, antiparasitärer, antimykotischer, antibakterieller und antiviraler Wirkung bei Mensch und Tier.

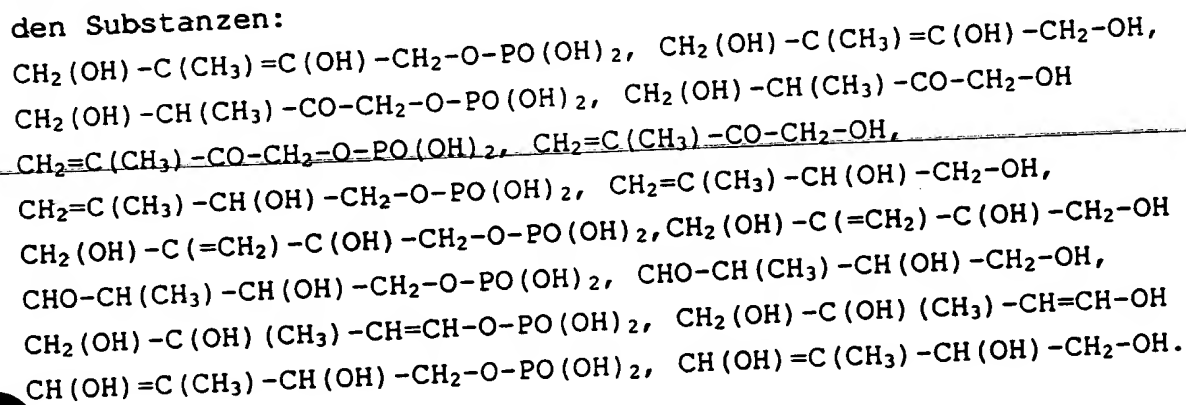
Der Biosyntheseweg zur Bildung von Isoprenoiden über den klassischen Acetat/ Mevalonat-Weg und einen alternativen, Mevalonat-unabhängigen Biosyntheseweg, den Desoxy-D-xylulose-Phosphat-Weg, ist bereits bekannt (Rohmer, M., Knani, M., Simonin, P., Sutter, B., and Sahm, H. (1993): Biochem. J. 295: 517-524).

Es ist aber nicht bekannt, wie und über welche Wege in Viren, Eukaryonten und Prokaryonten eine Änderung der Isoprenoidkonzentration über den Desoxy-D-xylulose-Phosphat-Weg erreicht werden kann. In Fig. 1 ist dieser Biosyntheseweg dargestellt.

Es werden daher DNA-Sequenzen zur Verfügung gestellt, die für die 1-Desoxy-D-xylulose-5-phosphat-Synthase (DOXP-Synthase), : 1-Desoxy-D-xylulose-5-phosphatreduktoisomerase (DOXP-Reduktoisomerase) oder das essentielle gcpE-Protein kodieren. Alle drei Gene und Enzyme sind an der Isoprenoid-Biosynthese beteiligt.

Das gcpE-Protein hat zusätzlich noch eine Kinasefunktion und katalysiert die Phosphorylierung eines Zuckers oder eines Phosphorzuckers oder einer Vorstufe der Isoprenoidbiosynthese, insbesondere die Phosphorylierung von 2-C-Methyl-D-erythritol, 2-C-Methyl-D-erythritol-phosphat, insbesondere 2-C-Methyl-D-

erythritol-4-phosphat, 2-C-Methyl-D-erythrose, 2-C-Methyl-D-erythrose-phosphat, insbesondere 2-C-Methyl-D-erythrose-4-phosphat. In der Vorstufe der Isoprenoidsynthese katalysiert das gcpE-Protein insbesondere die Phosphorylierung der folgenden Substanzen:



Die DOXP-Synthase katalysiert die Kondensation von Pyruvat und Glyceraldehyd-3-phosphat zu 1-Deoxy-D-xylulose-5-phosphat und die DOXP-Reduktoisomerase katalysiert die Umwandlung von 1-Deoxy-D-xylulose-5-phosphat zu 2-C-Methyl-D-erythritol-4-phosphat. (siehe Fig. 1).

Die Erfindung betrifft die folgenden DNA-Sequenzen: DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO:2 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:2, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren,

DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO:4 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:4, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren,

sowie DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO:6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:6, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch

andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

Die Gene und ihre Genprodukte (Polypeptide) sind im Sequenzprotokoll mit ihrer Primärstruktur aufgeführt und haben folgende Zuordnung:

SEQ ID NO:1: 1-Desoxy-D-xylulose-5-phosphatreduktoisomerase-Gen

SEQ ID NO:2: 1-Desoxy-D-xylulose-5-phosphatreduktoisomerase

SEQ ID NO:3: 1-Desoxy-D-xylulose-5-phosphat-Synthase-Gen

SEQ ID NO:4: 1-Desoxy-D-xylulose-5-phosphat-Synthase

SEQ ID NO:5: gcpE-Gen

SEQ ID NO:6: gcpE-Protein.

Die DNA-Sequenzen stammen alle aus Plasmodium falciparum.

Außer den im Sequenzprotokoll genannten DNA-Sequenzen sind auch solche geeignet, die infolge der Degeneration des genetischen Codes eine andere DNA-Sequenz besitzen, jedoch für das gleiche Polypeptid oder für ein Analoges oder Derivat des Polypeptids gemäß, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

Die erfindungsgemäßen Sequenzen eignen sich für die Expression von Genen in Viren, Eukaryonten und Prokaryonten, die für die Isoprenoid-Biosynthese des 1-Desoxy-D-xylulose-Wegs verantwortlich sind.

Erfindungsgemäß gehören zu den Eukaryonten oder eukaryontischen Zellen tierischen Zellen, Pflanzenzellen, Algen, Hefen, Pilze und zu den Prokaryonten oder prokaryontischen Bakterien Archaeobakterien und Eubakterien.

Bei Integration einer DNA-Sequenz in ein Genom, auf der eine der oben angegebenen DNA-Sequenzen lokalisiert ist, wird die Expression der oben beschriebenen Gene in Viren, Eukaryonten und Prokaryonten ermöglicht. Die erfindungsgemäß transformierten Viren, Eukaryonten und Prokaryonten werden in an sich be-

kannter Weise gezüchtet und das währenddessen gebildete Isoprenoid isoliert und gegebenenfalls gereinigt. Nicht alle Isoprenoide müssen isoliert werden, da die Isoprenoide in einigen Fällen direkt in die Raumluft abgegeben werden.

Die Erfindung betrifft ferner ein Verfahren zur Herstellung von transgenen Viren, Eukaryonten und Prokaryonten zur Veränderung des Isoprenoid-Gehaltes, das die folgenden Schritte enthält.

- a) Herstellung einer DNA-Sequenz mit folgenden Teilsequenzen
  - i) Promotor, der in Viren, Eukaryonten und Prokaryonten aktiv ist und die Bildung einer RNA im vorgesehenen Zielgewebe oder den Zielzellen sicherstellt,
  - ii) DNA-Sequenz, die für ein Polypeptid mit der in SEQ ID NO:2,4 oder 6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:2,4 oder 6,
  - iii) 3'-nichttranslatierte Sequenz, die in Viren, Eukaryonten und Prokaryonten zur Addition von Poly-A Resten an das 3'-Ende der RNA führt,
- b) Transfer und Einbau der DNA-Sequenz in das Genom von Viren, prokaryontischen oder eukaryontischen Zellen mit oder ohne Verwendung eines Vektors (z.B. Plasmid, virale DNA).

Aus ~~den~~ transformierten Pflanzenzellen können die intakten ganzen Pflanzen regeneriert werden.



Die für die Proteine kodierenden Sequenzen der Proteine mit den Nukleotidabfolgen Seq ID NO:1, Seq ID NO:3 und Seq ID NO: 5 können mit einem die Transkription in bestimmten Organen oder Zellen sicherstellenden Promotor versehen werden, der in sense-Orientierung (3'-Ende des Promotors zum 5'-Ende der kodierenden Sequenz) an die Sequenz, die das zu bildende Protein kodiert, gekoppelt ist. An das 3'-Ende der kodierenden Sequenz wird ein die Termination der mRNA-Synthese bestimmendes Terminationssignal angehängt. Um das zu exprimierende Protein in bestimmte subzelluläre Kompartimente, wie Chloroplasten, Amyloplasten, Mitochondrien, Vakuole, Cytosol oder Interzellularräume zu dirigieren, kann zwischen den Promotor und die ko-

dierende Sequenz noch eine für eine sogenannte Signalsequenz oder ein Transitpeptid kodierende Sequenz gesetzt werden. die Sequenz muß im gleichen Leserahmen wie die kodierende Sequenz des Proteins sein. Zur Vorbereitung der Einführung der erfindungsgemäßen DNA-Sequenzen in höhere Pflanzen sind eine große Anzahl von Klonierungsvektoren verfügbar, die ein Replikationssignal für E.coli und einen Marker beinhalten, der eine Selektion der transformierten Zellen erlaubt. Beispiele für Vektoren sind pBR 322, pUC-Serien, M13mp-Serien, pACYC 184, EMBL 3 usw. Je nach Einführungsmethode gewünschter Gene in die Pflanze können weitere DNA-Sequenzen erforderlich sein. Werden zum Beispiel für die Transformation der Pflanzenzelle das Ti- oder Ri-Plasmid verwendet, so muß mindestens eine rechte Begrenzung, häufig jedoch die rechte und die linke Begrenzung der Ti- und Ri-Plasmid T-DNA als Flankenbereich den einzuführenden Genen eingefügt werden. Die Verwendung von T-DNA für die Transformation von Pflanzenzellen ist intensiv untersucht und ausreichend in EP 120516; Hoekama, in: The Binary Plant Vector System, Offset-drukkerij Kanters B.V. Alblasterdam (1985), Chapter V; Fraley et al., Crit.Rev.Plant Sci. 4,1-46 und An et al. (1985) EMBO J. 4, 277-287 beschrieben worden. Ist die eingeführte DNA einmal im Genom integriert, so ist sie in der Regel stabil und bleibt auch in den Nachkommen der ursprünglich transformierten Zellen erhalten. Sie erhält normalerweise einen Selektionsmarker, der den transformierten Pflanzenzellen Resistenz gegenüber einem Biozid oder einem Antibiotikum, wie Kanamycin, G 418, Bleomycin, Hygromycin oder Phosphinotricin u.a. vermittelt. Der individuell verwendete Marker sollte daher die selektion transformierter Zellen gegenüber Zellen, denen die eingefügte DNA fehlt, gestatten.

Für die Einführung von DNA in eine Pflanze stehen viele Techniken zur Verfügung. Diese Techniken umfassen die Transformation mit Hilfe von Agrobakterien, z.B. Agrobacterium tumefaciens, die Fusion von Protoplasten, die Mikroinjektion von DNA, die Elektroporation, sowie ballistische Methoden und die Virusinfektion. Aus dem transformierten Pflanzenmaterial können dann im geeigneten Medium, welches Antibiotika oder Biozide zur Selektion enthalten kann, wieder ganze Pflanzen regeneriert werden.

Bei der Injektion und Elektroporation sind an sich keine speziellen Anforderungen an die Plasmide gestellt. Sollen aber aus derartig transformierten Zellen ganze Pflanzen regeneriert werden, ist die Anwesenheit eines selektierbaren Markergens notwendig. Die transformierten Zellen wachsen innerhalb der Pflanzen in der üblichen Weise (McCormick et al. (1986), Plant Cell Reports 5, 81-84). Die Pflanzen können normal angezogen werden und mit Pflanzen, die die gleiche transformierte Erbanlage oder andere Erbanlagen haben, gekreuzt werden. Die daraus entstehenden Individuen haben die entsprechenden phänotypischen Eigenschaften.



Weiterhin sind Gegenstand der Erfindung Expressionsvektoren, die eine oder mehrere der erfindungsgemäßen DNA-Sequenzen enthalten. Solche Expressionsvektoren erhält man, indem man die erfindungsgemäßen DNA-Sequenzen mit geeigneten funktionellen Regulationssignalen versieht. Solche Regulationssignale sind DNA-Sequenzen, die für die Expression verantwortlich sind, beispielsweise Promotoren, Operatoren, Enhancer, ribosomale Bindungsstellen, und die vom Wirtsorganismus erkannt werden.

Gegebenenfalls können noch weitere Regulationssignale, die beispielsweise Replikation oder Rekombination der rekombinanten DNA im Wirtsorganismus steuern, Bestandteil des Expressionsvektors sein.

Ebenso gehören die mit den erfindungsgemäßen DNA-Sequenzen oder Expressionsvektoren transformierten Wirtsorganismen zum Gegenstand der Erfindung.

Für die Expression der erfindungsgemäßen Enzyme eignen sich besonders solche Wirtszellen und Organismen, die keine intrinsischen Enzyme mit der Funktion der DOXP-Synthase, der DOXP-Reduktoisomerase oder des gcpE-Proteins aufweisen. Dies trifft für Archaeobakterien, Tiere, Pilze, Schleimpilze und einige Eubakterien zu. Durch das Fehlen dieser intrinsischen Enzymaktivitäten wird die Detektion und Aufreinigung der rekombinanten Enzyme wesentlich erleichtert. Auch wird es erst dadurch mög-

lich, mit geringem Aufwand die Aktivität und insbesondere die Hemmung der Aktivität der erfindungsgemäßen rekombinanten Enzyme durch verschiedenen Chemikalien und Pharmaka in Rohextrakten aus den Wirtszellen zu messen.

Die Expression der erfindungsgemäßen Enzyme erfolgt vorteilhafterweise dann in eukaryontischen Zellen, wenn posttranslatorische Modifikationen und eine native Faltung der Polypeptidkette erreicht werden soll. Außerdem wird in Abhängigkeit vom Expressionssystem bei der Expression genomischer DNA-Sequenzen erreicht, daß Introns durch Spleißen der DNA beseitigt und die Enzyme in der für die Parasiten charakteristischen Polypeptidsequenz produziert werden. Für Introns codierende Sequenzen können auch durch rekombinante DNA-Technologie aus den zu exprimierenden DNA-Sequenzen beseitigt oder experimentell eingefügt werden.

Die Isolierung des Proteins kann aus der Wirtszelle oder dem Kulturüberstand der Wirtszelle nach dem Fachmann bekannten Verfahren erfolgen. Es kann auch eine in vitro Reaktivierung der Enzyme erforderlich sein.

Zur Erleichterung der Aufreinigung können die erfindungsgemäßen Enzyme oder Teilsequenzen der Enzyme als Fusionsprotein mit verschiedenen Peptidketten exprimiert werden. Dazu eignen sich besonders Oligo-Histidin-Sequenzen und Sequenzen, die von der Glutathion-S-Transferase, Thioredoxin oder Calmodulin-bindenden Peptiden abgeleitet sind. Fusionen mit Thioredoxin-abgeleiteten Sequenzen eignen sich besonders für prokaryontische Expression, da dadurch die Löslichkeit der rekombinanten Enzyme erhöht wird.

Weiterhin können die erfindungsgemäßen Enzyme oder Teilsequenzen der Enzyme als Fusionsprotein mit solchen, dem Fachmann bekannten, Peptidketten exprimiert werden, daß die rekombinanten



Enzyme in das extrazelluläre Milieu oder in bestimmte Kompartimente der Wirtszellen transportiert werden. Dadurch kann sowohl die Aufreinigung, als auch die Untersuchung der biologischen Aktivität der Enzyme erleichtert werden.

Bei der Expression der erfindungsgemäßen Enzyme kann es sich als zweckmäßig erweisen, einzelne Codone zu verändern. Dabei ist der gezielte Austausch von Basen in der kodierenden Region auch sinnvoll, wenn die genutzten Codone in den Parasiten abweichend sind von der Codonnutzung im heterologen Expressionssystem, um eine optimale Synthese des Proteins zu gewährleisten. Zudem sind oft Deletionen von nicht-translatierten 5'-bzw. 3'-Abschnitten sinnvoll, beispielsweise wenn mehrere destabilisierende Sequenzmotive ATTTA im 3'-Bereich der DNA vorliegen. Dann sollten diese bei der bevorzugten Expression in Eukaryonten deletiert werden. Veränderungen dieser Art sind Deletionen, Additionen oder Austausch von Basen und ebenfalls Gegenstand der vorliegenden Erfindung.

Weiterhin können die erfindungsgemäßen Enzyme unter standardisierten Bedingungen durch dem Fachmann bekannte Techniken durch in vitro-Translation gewonnen werden. Dafür geeignete Systeme sind Kaninchen-Reticulozyten- und Weizenkeimextrakte und Bakterienlysate. Auch kann in vitro transskribierte mRNA in Xenopus-Oocyten translatiert werden.

Durch chemische Synthese können Oligo- und Polypeptide hergestellt werden, der Sequenzen aus der Peptidsequenz der erfindungsgemäßen Enzyme abgeleitet sind. Bei geeigneter Wahl der Sequenzen besitzen derartige Peptide Eigenschaften, die für die vollständigen erfindungsgemäßen Enzyme charakteristisch sind. Derartige Peptide können in großen Mengen hergestellt werden und eignen sich besonders für Studien über die Kinetik der Enzymaktivität, die Regulation der Enzymaktivität, die dreidimensionale Struktur der Enzyme, die Hemmung der Enzymaktivität

97 05.11.99

durch verschiedenen Chemikalien und Pharmaka und die Bindungs-  
geometrie und Bindungsaffinität verschiedener Liganden.

Vorzugsweise wird zur rekombinanten Herstellung der erfindungs-  
gemäßen Enzyme eine DNA mit den Nukleotiden aus den Sequenzen  
SEQ ID NO: 1, 3 und 5 verwendet.

---

~~Wie in dem dieser Anmeldung zugrundeliegenden Prioritätsdoku-~~  
ment beschrieben, hat sich herausgestellt, daß in vielen Para-  
siten, Bakterien, Viren und Pilzen dieser der Desoxy-D-  
xylulose-Phosphat-Stoffwechselweg ebenfalls vorliegt.

Die Erfindung umfaßt daher außerdem ein Verfahren zum Screening  
einer Verbindung. Gemäß diesem Verfahren wird ein Wirtsorganis-  
mus, der einen rekombinanten Expressionsvektor enthält, wobei  
der Vektor zumindest einen Teil der Oligonukleotidsequenz gemäß  
SEQ ID NO: 1, SEQ ID NO: 3 oder SEQ ID NO: 5 oder Varianten oder  
Homologe dieser aufweist, und außerdem eine Verbindung, von der  
vermutet wird, daß sie eine antimikrobielle, antiparasitäre,  
antibakterielle, antivirale und antimykotische Wirkung bei  
Mensch und Tier oder eine antimikrobielle, antivirale, bakteri-  
zide, herbizide oder fungizide Wirkung bei Pflanzen hat, be-  
reitgestellt. Anschließend wird der Wirtsorganismus mit der  
Verbindung in Kontakt gebracht und die Wirksamkeit der Verbin-  
dung bestimmt.

Ein weiterer Gegenstand dieser Erfindung sind Methoden zur Be-  
stimmung der enzymatischen Aktivität des gcpE-Proteins. Diese  
kann nach den bekannten Anleitungen bestimmt werden. Hierbei  
wird die Phosphorylierung eines Zuckers oder eines Phosphorzuk-  
kers oder einer Vorstufe der Isoprenoidbiosynthese, insbeson-  
dere die Phosphorylierung von 2-C-Methyl-D-erythritol, 2-C-  
Methyl-D-erythritol-phosphat, insbesondere 2-C-Methyl-D-  
erythritol-4-phosphat, 2-C-Methyl-D-erythrose, 2-C-Methyl-D-  
erythrose-phosphat, insbesondere 2-C-Methyl-D-erythrose-4-  
phosphat, detektiert. Ein weiterer Gegenstand dieser Erfindung  
ist die Verwendung dieser Meßverfahren zur Ermittlung von Stof-  
fen, die die Aktivität der jeweiligen Enzyme inhibieren.

PI 05.11.99

Analog erfolgt die Bestimmung der Aktivitäten von DOXP-Synthase und DOXP-Reduktoisomerase. Für die Bestimmung der DOXP-Synthase-Aktivität eignen sich auch fluorimetrische Verfahren, wie von Querol et al. beschrieben (Querol et al. Abstracts 4<sup>th</sup> european symposium on plant isoprenoids, Barcelona 21-23 April 1999).

---

# Sequenzprotokoll

Anzahl der Sequenzen: 6

(1) ANGABEN ZU SEQUENZ ID NO: 1

Plasmodium falciparum 1-Desoxy-D-xylulose-5-phosphatredukto-  
isomerase(dxr)gen

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 1467 BASENPAARE

(B) ART: Nukleotidsequenz

(C) STAMM: HB3

(ii) ART DES MOLEKÜLS: DNA

(iv) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Plasmodium falciparum

(ix) MERKMAL

(A) NAME/SCHLÜSSEL: mRNA

(B) LAGE:1...1467

GEN=dxr

PRODUKT=1-Desoxy-D-xylulose-5-phosphatreduktoisomerase

MERKMAL

(A) NAME/SCHLÜSSEL: Gen

(B) LAGE:1...1467

GEN=dxr

(ix) MERKMAL

(A) NAME/SCHLÜSSEL: CDS

(B) LAGE:1...1467

GEN=dxr

FUNKTION: bei der Isopentenylidiphosphatbiosynthese betei-  
ligt

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:1

ATG AAG AAA TAT ATT TAT ATA TAT TTT TTC TTC ATC ACA ATA ACT ATT	48
Met Lys Lys Tyr Ile Tyr Ile Tyr Phe Phe Phe Ile Thr Ile Thr Ile	
5 10 15	
AAT GAT TTA GTA ATA AAT AAT ACA TCA AAA TGT GTT TCC ATT GAA AGA	96
Asn Asp Leu Val Ile Asn Asn Thr Ser Lys Cys Val Ser Ile Glu Arg	
20 25 30	
AGA AAA AAT AAC GCA TAT ATA AAT TAT GGT ATA GGA TAT AAT GGA CCA	144
Arg Lys Asn Asn Ala Tyr Ile Asn Tyr Glu Ile Glu Tyr Asn Glu Pro	
35 40 45	

GAT	AAT	AAA	ATA	ACA	AAG	AGT	AGA	AGA	TGT	AAA	AGA	ATA	AAG	TTA	TGC	192
Asp	Asn	Lys	Ile	Thr	Lys	Ser	Arg	Arg	Cys	Lys	Arg	Ile	Lys	Leu	Cys	
	50					55				60						
AAA	AAG	GAT	TTA	ATA	GAT	ATT	GGT	GCA	ATA	AAG	AAA	CCA	ATT	AAT	GTA	240
Lys	Lys	Asp	Leu	Ile	Asp	Ile	Glu	Ala	Ile	Lys	Lys	Pro	Ile	Asn	Val	
65					70					75					80	
GCA	ATT	TTT	GGA	AGT	ACT	GGT	AGT	ATA	GGT	ACG	AAT	GCT	TTA	AAT	ATA	288
Ala	Ile	Phe	Glu	Ser	Thr	Glu	Ser	Ile	Glu	Thr	Asn	Ala	Leu	Asn	Ile	
				85					90					95		
ATA	AGG	GAG	TGT	AAT	AAA	ATT	GAA	AAT	GTT	TTT	AAT	GTT	AAA	GCA	TTG	336
Ile	Arg	Glu	Cys	Asn	Lys	Ile	Glu	Asn	Val	Phe	Asn	Val	Lys	Ala	Leu	
			100					105				110				
TAT	GTG	AAT	AAG	AGT	GTG	AAT	GAA	TTA	TAT	GAA	CAA	GCT	AGA	GAA	TTT	384
Tyr	Val	Asn	Lys	Ser	Val	Asn	Glu	Leu	Tyr	Glu	Gln	Ala	Arg	Glu	Phe	
		115					120					125				
TTA	CCA	GAA	TAT	TTG	TGT	ATA	CAT	GAT	AAA	AGT	GTA	TAT	GAA	GAA	TTA	432
Leu	Pro	Glu	Tyr	Leu	Cys	Ile	His	Asp	Lys	Ser	Val	Tyr	Glu	Glu	Leu	
130						135					140					
AAA	GAA	CTG	GTA	AAA	AAT	ATA	AAA	GAT	TAT	AAA	CCT	ATA	ATA	TTG	TGT	480
Lys	Glu	Leu	Val	Lys	Asn	Ile	Lys	Asp	Tyr	Lys	Pro	Ile	Ile	Leu	Cys	
145					150					155				160		
GGT	GAT	GAA	GGG	ATG	AAA	GAA	ATA	TGT	AGT	AGT	AAT	AGT	ATA	GAT	AAA	528
Glu	Asp	Glu	Glu	Met	Lys	Glu	Ile	Cys	Ser	Ser	Asn	Ser	Ile	Asp	Lys	
				165				170						175		
ATA	GTT	ATT	GGT	ATT	GAT	TCT	TTT	CAA	GGA	TTA	TAT	TCT	ACT	ATG	TAT	576
Ile	Val	Ile	Glu	Ile	Asp	Ser	Phe	Gln	Glu	Leu	Tyr	Ser	Thr	Met	Tyr	
			180					185					190			
GCA	ATT	ATG	AAT	AAT	AAA	ATA	GTT	GCG	TTA	GCT	AAT	AAA	GAA	TCC	ATT	624
Ala	Ile	Met	Asn	Asn	Lys	Ile	Val	Ala	Leu	Ala	Asn	Lys	Glu	Ser	Ile	
		195					200					205				
GTC	TCT	GCT	GGT	TTC	TTT	TTA	AAG	AAA	TTA	TTA	AAT	ATT	CAT	AAA	AAT	672
Val	Ser	Ala	Glu	Phe	Phe	Leu	Lys	Lys	Leu	Leu	Asn	Ile	His	Lys	Asn	
	210					215					220					
GCA	AAG	ATA	ATA	CCT	GTT	GAT	TCA	GAA	CAT	AGT	GCT	ATA	TTT	CAA	TGT	720
Ala	Lys	Ile	Ile	Pro	Val	Asp	Ser	Glu	His	Ser	Ala	Ile	Phe	Gln	Cys	
225				230				235						240		
TTA	GAT	AAT	AAT	AAG	GTA	TTA	AAA	ACA	AAA	TGT	TTA	CAA	GAC	AAT	TTT	768
Leu	Asp	Asn	Asn	Lys	Val	Leu	Lys	Thr	Lys	Cys	Leu	Gln	Asp	Asn	Phe	
				245				250						255		
TCT	AAA	ATT	AAC	AAT	ATA	AAT	AAA	ATA	TTT	TTA	TGT	TCA	TCT	GGA	GGT	816
Ser	Lys	Ile	Asn	Asn	Ile	Asn	Lys	Ile	Phe	Leu	Cys	Ser	Ser	Glu	Glu	
			260					265					270			
CCA	TTT	CAA	AAT	TTA	ACT	ATG	GAC	GAA	TTA	AAA	AAT	GTA	ACA	TCA	GAA	864
Pro	Phe	Gln	Asn	Leu	Thr	Met	Asp	Glu	Leu	Lys	Asn	Val	Thr	Ser	Glu	
		275					280					285				
AAT	GCT	TTA	AAG	CAT	CCT	AAA	TGG	AAA	ATG	GGT	AAG	AAA	ATA	ACT	ATA	912
Asn	Ala	Leu	Lys	His	Pro	Lys	Trp	Lys	Met	Glu	Lys	Lys	Ile	Thr	Ile	
	290					295				300						
GAT	TCT	GCA	ACT	ATG	ATG	AAT	AAA	GGT	TTA	GAG	GTT	ATA	GAA	ACC	CAT	960
Asp	Ser	Ala	Thr	Met	Met	Asn	Lys	Glu	Leu	Glu	Val	Ile	Glu	Thr	His	
305				310				315						320		
TTT	TTA	TTT	GAT	GTA	GAT	TAT	AAT	GAT	ATA	GAA	GTT	ATA	GTA	CAT	AAA	1008
Phe	Leu	Phe	Asp	Val	Asp	Tyr	Asn	Asp	Ile	Glu	Val	Ile	Val	His	Lys	
			325					330						335		
GAA	TGC	ATT	ATA	CAT	TCT	TGT	GTT	GAA	TTT	ATA	GAC	AAA	TCA	GTA	ATA	1056
Glu	Cys	Ile	Ile	His	Ser	Cys	Val	Glu	Phe	Ile	Asp	Lys	Ser	Val	Ile	
			340					345					350			

AGT	CAA	ATG	TAT	TAT	CCA	GAT	ATG	CAA	ATA	CCC	ATA	TTA	TAT	TCT	TTA	1104
Ser	Gln	Met	Tyr	Tyr	Pro	Asp	Met	Gln	Ile	Pro	Ile	Leu	Tyr	Ser	Leu	
		355					360					365				
ACA	TGG	CCT	GAT	AGA	ATA	AAA	ACA	AAT	TTA	AAA	CCT	TTA	GAT	TTG	GCT	1152
Thr	Trp	Pro	Asp	Arg	Ile	Lys	Thr	Asn	Leu	Lys	Pro	Leu	Asp	Leu	Ala	
	370					375				380						
CAG	GTT	TCA	ACT	CTT	ACA	TTT	CAT	AAA	CCT	TCT	TTA	GAA	CAT	TTC	CCG	1200
Gln	Val	Ser	Thr	Leu	Thr	Phe	His	Lys	Pro	Ser	Leu	Glu	His	Phe	Pro	
	385				390				395						400	
TGT	ATT	AAA	TTA	GCT	TAT	CAA	GCA	GGT	ATA	AAA	GGA	AAC	TTT	TAT	CCA	1248
Cys	Ile	Lys	Leu	Ala	Tyr	Gln	Ala	Glu	Ile	Lys	Glu	Asn	Phe	Tyr	Pro	
				405				410					415			
ACT	GTA	CTA	AAT	GCG	TCA	AAT	GAA	ATA	GCT	AAC	AAC	TTA	TTT	TTG	AAT	1296
Thr	Val	Leu	Asn	Ala	Ser	Asn	Glu	Ile	Ala	Asn	Asn	Leu	Phe	Leu	Asn	
			420					425					430			
AAT	AAA	ATT	AAA	TAT	TTT	GAT	ATT	TCC	TCT	ATA	ATA	TCG	CAA	GTT	CTT	1344
Asn	Lys	Ile	Lys	Tyr	Phe	Asp	Ile	Ser	Ser	Ile	Ile	Ser	Gln	Val	Leu	
		435					440					445				
GAA	TCT	TTC	AAT	TCT	CAA	AAG	GTT	TCG	GAA	AAT	AGT	GAA	GAT	TTA	ATG	1392
Glu	Ser	Phe	Asn	Ser	Gln	Lys	Val	Ser	Glu	Asn	Ser	Glu	Asp	Leu	Met	
	450					455				460						
AAG	CAA	ATT	CTA	CAA	ATA	CAT	TCT	TGG	GCC	AAA	GAT	AAA	GCT	ACC	GAT	1440
Lys	Gln	Ile	Leu	Gln	Ile	His	Ser	Trp	Ala	Lys	Asp	Lys	Ala	Thr	Asp	
	465				470				475						480	
ATA	TAC	AAC	AAA	CAT	AAT	TCT	TCA	TAG								1467
Ile	Tyr	Asn	Lys	His	Asn	Ser	Ser									
					485											

(2) Angaben zu Sequenz ID No: 2

(i) Sequenzkennzeichen:

(A) Länge: 488 Aminosäuren

(B) Art: Aminosäure

(ii) Art des Moleküls: Protein

(iii) Ursprüngliche Herkunft:

(A) ORGANISMUS: Plasmodium falciparum; (Apicomplexa)

STAMM: HB3

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:2

Met	Lys	Lys	Tyr	Ile	Tyr	Ile	Tyr	Phe	Phe	Phe	Ile	Thr	Ile	Thr	Ile
1				5					10				15		
Asn	Asp	Leu	Val	Ile	Asn	Asn	Thr	Ser	Lys	Cys	Val	Ser	Ile	Glu	Arg
		20					25					30			
Arg	Lys	Asn	Asn	Ala	Tyr	Ile	Asn	Tyr	Gly	Ile	Gly	Tyr	Asn	Gly	Pro
	35					40					45				
Asp	Asn	Lys	Ile	Thr	Lys	Ser	Arg	Arg	Cys	Lys	Arg	Ile	Lys	Leu	Cys
	50				55				60						
Lys	Lys	Asp	Leu	Ile	Asp	Ile	Gly	Ala	Ile	Lys	Lys	Pro	Ile	Asn	Val
65				70					75					80	

Ala	Ile	Phe	Gly	Ser	Thr	Gly	Ser	Ile	Gly	Thr	Asn	Ala	Leu	Asn	Ile
				85					90					95	
Ile	Arg	Glu	Cys	Asn	Lys	Ile	Glu	Asn	Val	Phe	Asn	Val	Lys	Ala	Leu
		100					105						110		
Tyr	Val	Asn	Lys	Ser	Val	Asn	Glu	Leu	Tyr	Glu	Gln	Ala	Arg	Glu	Phe
		115					120					125			
Leu	Pro	Glu	Tyr	Leu	Cys	Ile	His	Asp	Lys	Ser	Val	Tyr	Glu	Glu	Leu
	130					135					140				
Lys	Glu	Leu	Val	Lys	Asn	Ile	Lys	Asp	Tyr	Lys	Pro	Ile	Ile	Leu	Cys
145					150					155					160
<del>Gly</del>	<del>Asp</del>	<del>Glu</del>	<del>Gly</del>	<del>Met</del>	<del>Lys</del>	<del>Glu</del>	<del>Ile</del>	<del>Cys</del>	<del>Ser</del>	<del>Ser</del>	<del>Asn</del>	<del>Ser</del>	<del>Ile</del>	<del>Asp</del>	<del>Lys</del>
Ile	Val	Ile	Gly	Ile	Asp	Ser	Phe	Gln	Gly	Leu	Tyr	Ser	Thr	Met	Tyr
Ala	Ile	Met	Asn	Asn	Lys	Ile	Val	Ala	Leu	Ala	Asn	Lys	Glu	Ser	Ile
Val	Ser	Ala	Gly	Phe	Phe	Leu	Lys	Lys	Leu	Leu	Asn	Ile	His	Lys	Asn
Ala	Lys	Ile	Ile	Pro	Val	Asp	Ser	Glu	His	Ser	Ala	Ile	Phe	Gln	Cys
Leu	Asp	Asn	Asn	Lys	Val	Leu	Lys	Thr	Lys	Cys	Leu	Gln	Asp	Asn	Phe
Ser	Lys	Ile	Asn	Asn	Ile	Asn	Lys	Ile	Phe	Leu	Cys	Ser	Ser	Gly	Gly
Pro	Phe	Gln	Asn	Leu	Thr	Met	Asp	Glu	Leu	Lys	Asn	Val	Thr	Ser	Glu
Asn	Ala	Leu	Lys	His	Pro	Lys	Trp	Lys	Met	Gly	Lys	Lys	Ile	Thr	Ile
Asp	Ser	Ala	Thr	Met	Met	Asn	Lys	Gly	Leu	Glu	Val	Ile	Glu	Thr	His
Phe	Leu	Phe	Asp	Val	Asp	Tyr	Asn	Asp	Ile	Glu	Val	Ile	Val	His	Lys
Glu	Cys	Ile	Ile	His	Ser	Cys	Val	Glu	Phe	Ile	Asp	Lys	Ser	Val	Ile
Ser	Gln	Met	Tyr	Tyr	Pro	Asp	Met	Gln	Ile	Pro	Ile	Leu	Tyr	Ser	Leu
Thr	Trp	Pro	Asp	Arg	Ile	Lys	Thr	Asn	Leu	Lys	Pro	Leu	Asp	Leu	Ala
Gln	Val	Ser	Thr	Leu	Thr	Phe	His	Lys	Pro	Ser	Leu	Glu	His	Phe	Pro
Cys	Ile	Lys	Leu	Ala	Tyr	Gln	Ala	Gly	Ile	Lys	Gly	Asn	Phe	Tyr	Pro
Thr	Val	Leu	Asn	Ala	Ser	Asn	Glu	Ile	Ala	Asn	Asn	Leu	Phe	Leu	Asn
Asn	Lys	Ile	Lys	Tyr	Phe	Asp	Ile	Ser	Ser	Ile	Ile	Ser	Gln	Val	Leu
Glu	Ser	Phe	Asn	Ser	Gln	Lys	Val	Ser	Glu	Asn	Ser	Glu	Asp	Leu	Met
Lys	Gln	Ile	Leu	Gln	Ile	His	Ser	Trp	Ala	Lys	Asp	Lys	Ala	Thr	Asp
Ile	Tyr	Asn	Lys	His	Asn	Ser	Ser								

(3) ANGABEN ZU SEQUENZ ID NR: 3  
Plasmodium Falciparum 1-Desoxy-D-Xylulose-5-phosphatsynthase(dxsgen

(iii) SEQUENZKENNZEICHEN:

(A) LÄNGE: 3872 BASENPAARE

(B) ART: Nukleotidsequenz

(C) STAMM: HB3

(iv) ART DES MOLEKÜLS: DNA

(v) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Plasmodium falciparum

(ix) MERKMAL

(A) NAME/SCHLÜSSEL: mRNA

GEN=dxs

PRODUKT=1-Desoxy-D-xylulose-5-phosphatsynthase

(ix) MERKMAL

(A) NAME/SCHLÜSSEL: Gen

(B) LÄNGE: 1..3872

GEN=dxs

(ix) MERKMAL

(A) NAME/SCHLÜSSEL: CDS

GEN=dxs

FUNKTION: bei der Isopentenylidiphosphatbiosynthese beteiligt

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:3

```

GTAATATAC GTATAATATA TATATAATAT ATTCTTACGT ATGTATCATT TATGAATCAT 60
AATAATATTC TAAATTTACC TTCCGTTTTT GCTCGATCTT CTCATTTTCG TTTCAGCTTT 120
TATCA ATG ATT TTT AAT TAT GTG TTT TTT AAG AAC TTT GTA CCA GTT GTT 170
      Met Ile Phe Asn Tyr Val Phe Phe Lys Asn Phe Val Pro Val Val
        1             5             10             15
CTA TAC ATT CTC CTT ATA ATA TAT ATT AAC TTA AAT GGC ATG AAT AAT 218
Leu Tyr Ile Leu Leu Ile Ile Tyr Ile Asn Leu Asn Gly Met Asn Asn
              20             25             30
AAA AAT CAA ATA AAA ACA GAA AAA ATT TAT ATA AAG AAA TTG AAT AGG 266
Lys Asn Gln Ile Lys Thr Glu Lys Ile Tyr Ile Lys Lys Leu Asn Arg
              35             40             45
TTG TCA AGG AAA AAT TCG TTA TGT AGT TCT AAA AAT AAA ATA GCA TGC 314
Leu Ser Arg Lys Asn Ser Leu Cys Ser Ser Lys Asn Lys Ile Ala Cys
              50             55             60
TTG TTC GAT ATA GGA AAT GAT GAT AAT AGA AAT ACG ACA TAT GGC TAT 362
Leu Phe Asp Ile Gly Asn Asp Asp Asn Arg Asn Thr Thr Tyr Gly Tyr
        65             70             75

```



AAT	GTG	AAT	GTT	AAA	AAT	GAT	GAT	ATT	AAT	TCC	TTA	CTA	AAA	AAT	AAT	410
Asn	Val	Asn	Val	Lys	Asn	Asp	Asp	Ile	Asn	Ser	Leu	Leu	Lys	Asn	Asn	
80					85					90					95	
TAT	AGT	AAT	AAA	TTG	TAC	ATG	GAT	AAG	AGG	AAA	AAT	ATT	AAT	AAT	GTA	458
Tyr	Ser	Asn	Lys	Leu	Tyr	Met	Asp	Lys	Arg	Lys	Asn	Ile	Asn	Asn	Val	
				100					105						110	
ATT	AGT	ACT	AAT	AAA	ATA	TCT	GGG	TCC	ATT	TCA	AAT	ATT	TGT	AGT	AGA	506
Ile	Ser	Thr	Asn	Lys	Ile	Ser	Gly	Ser	Ile	Ser	Asn	Ile	Cys	Ser	Arg	
			115					120					125			
AAT	CAA	AAA	GAA	AAT	GAA	CAA	AAA	AGA	AAT	AAA	CAA	AGA	TGT	TTA	ACT	554
Asn	Gln	Lys	Glu	Asn	Glu	Gln	Lys	Arg	Asn	Lys	Gln	Arg	Cys	Leu	Thr	
		130					135					140				
CAA	TGT	CAC	ACT	TAT	AAT	ATG	TCA	CAT	GAA	CAG	GAC	AAA	CTA	GCT	AAT	602
Gln	Cys	His	Thr	Tyr	Asn	Met	Ser	His	Glu	Gln	Asp	Lys	Leu	Ala	Asn	
	145					150					155					
GAT	AAT	AAT	AGG	AAT	AAT	AAA	AAG	AAT	TTT	AAT	TTA	TTA	TTT	ATA	AAT	650
Asp	Asn	Asn	Arg	Asn	Asn	Lys	Lys	Asn	Phe	Asn	Leu	Leu	Phe	Ile	Asn	
160					165					170					175	
TAT	TTT	AAT	TTG	AAA	CGA	ATG	AAA	AAT	TCT	CTT	CTA	AAT	AAA	GAC	AAT	698
Tyr	Phe	Asn	Leu	Lys	Arg	Met	Lys	Asn	Ser	Leu	Leu	Asn	Lys	Asp	Asn	
				180					185					190		
TTC	TTT	TAC	TGT	AAA	GAA	AAA	AAA	TTG	TCA	TTT	CTG	CAT	AAG	GCC	TAT	746
Phe	Phe	Tyr	Cys	Lys	Glu	Lys	Lys	Leu	Ser	Phe	Leu	His	Lys	Ala	Tyr	
			195					200					205			
AAA	AAA	AAA	AAT	TGC	ACT	TTT	CAA	AAT	TAT	AGT	TTA	AAA	AGA	AAA	TCT	794
Lys	Lys	Lys	Asn	Cys	Thr	Phe	Gln	Asn	Tyr	Ser	Leu	Lys	Arg	Lys	Ser	
		210					215					220				
AAT	CGT	GAT	TCA	CAT	AAA	TTG	TTT	TCT	GGA	GAA	TTT	GAC	GAT	TAT	ACA	842
Asn	Arg	Asp	Ser	His	Lys	Leu	Phe	Ser	Gly	Glu	Phe	Asp	Asp	Tyr	Thr	
	225					230					235					
AAT	AAT	AAT	GCT	TTA	TAT	GAA	TCC	GAA	AAA	AAA	GAA	TAC	ATT	ACA	CTA	890
Asn	Asn	Asn	Ala	Leu	Tyr	Glu	Ser	Glu	Lys	Lys	Glu	Tyr	Ile	Thr	Leu	
240					245					250					255	
AAT	AAT	AAT	AAT	AAA	AAT	AAT	AAT	AAT	AAA	AAT	AAT	GAT	AAT	AAA	AAT	938
Asn	Asn	Asn	Asn	Lys	Asn	Asn	Asn	Asn	Lys	Asn	Asn	Asp	Asn	Lys	Asn	
				260					265					270		
AAT	GAT	AAT	AAT	GAT	TAT	AAT	AAT	AAT	AAT	AGT	TGT	AAT	AAT	TTA	GGA	986
Asn	Asp	Asn	Asn	Asp	Tyr	Asn	Asn	Asn	Asn	Ser	Cys	Asn	Asn	Leu	Gly	
			275					280					285			
GAG	AGA	TCC	AAT	CAT	TAT	GAT	AAT	TAT	GGT	GGA	GAT	AAT	AAT	AAT	CCA	1034
Glu	Arg	Ser	Asn	His	Tyr	Asp	Asn	Tyr	Gly	Gly	Asp	Asn	Asn	Asn	Pro	
		290					295					300				
TGT	AAT	AAT	AAT	AAT	GAC	AAA	TAT	GAT	ATA	GGA	AAA	TAT	TTC	AAA	CAG	1082
Cys	Asn	Asn	Asn	Asn	Asp	Lys	Tyr	Asp	Ile	Gly	Lys	Tyr	Phe	Lys	Gln	
		305				310					315					
ATT	AAT	ACC	TTT	ATT	AAT	ATT	GAT	GAA	TAT	AAA	ACT	ATA	TAT	GGT	GAT	1130
Ile	Asn	Thr	Phe	Ile	Asn	Ile	Asp	Glu	Tyr	Lys	Thr	Ile	Tyr	Gly	Asp	
320					325					330					335	
GAA	ATA	TAT	AAA	GAA	ATA	TAT	GAA	CTA	TAT	GTA	GAA	AGA	AAT	ATT	CCT	1178
Glu	Ile	Tyr	Lys	Glu	Ile	Tyr	Glu	Leu	Tyr	Val	Glu	Arg	Asn	Ile	Pro	
			340					345						350		
GAA	TAT	TAT	GAA	CGA	AAA	TAT	TTT	TCA	GAA	GAT	ATT	AAA	AAG	AGT	GTC	1226
Glu	Tyr	Tyr	Glu	Arg	Lys	Tyr	Phe	Ser	Glu	Asp	Ile	Lys	Lys	Ser	Val	
			355					360					365			
CTA	TTT	GAT	ATA	GAT	AAA	TAT	AAT	GAT	GTC	GAA	TTT	GAA	AAA	GCT	ATA	1274
Leu	Phe	Asp	Ile	Asp	Lys	Tyr	Asn	Asp	Val	Glu	Phe	Glu	Lys	Ala	Ile	
		370					375					380				

AAA Lys	GAA Glu	GAA Glu	TTT Phe	ATA Ile	AAT Asn	AAT Asn	GGA Gly	GTT Val	TAT Tyr	ATT Ile	AAT Asn	AAT Asn	ATA Ile	GAT Asp	AAT Asn	1322
ACA Thr	TAT Tyr	TAT Tyr	AAA Lys	AAA Lys	GAA Glu	AAT Asn	ATT Ile	TTA Leu	ATA Ile	ATG Met	AAA Lys	AAG Lys	ATA Ile	TTA Leu	CAT His	1370
TAT Tyr	TTC Phe	CCA Pro	TTA Leu	TTA Leu	AAA Lys	TTA Leu	ATT Ile	AAT Asn	AAT Asn	CCA Pro	TCA Ser	GAT Asp	TTA Leu	AAA Lys	AAG Lys	1418
TTA Leu	AAA Lys	AAA Lys	CAA Gln	TAT Tyr	TTA Leu	CCT Pro	TTA Leu	TTA Leu	GCA Ala	CAT His	GAA Glu	TTA Leu	AAA Lys	ATA Ile	TTT Phe	1466
TTA Leu	TTT Phe	TTT Phe	ATT Ile	GTA Val	AAT Asn	ATA Ile	ACA Thr	GGA Gly	GGT Gly	CAT His	TTT Phe	TCC Ser	TCT Ser	GTT Val	TTA Leu	1514
AGC Ser	TCT Ser	TTA Leu	GAA Glu	ATT Ile	CAA Gln	TTA Leu	TTA Leu	TTG Leu	TAT Tyr	ATT Ile	TTT Phe	AAT Asn	CAA Gln	CCA Pro		1562
TAT Tyr	GAT Asp	AAT Asn	GTT Val	ATA Ile	TAT Tyr	GAT Asp	ATA Ile	GGA Gly	CAT His	CAA Gln	GCA Ala	TAT Tyr	GTA Val	CAT His	AAG Lys	1610
ATA Ile	TTG Leu	ACC Thr	GGA Gly	AGA Arg	AAA Lys	CTA Leu	TTA Leu	TTT Phe	CTA Leu	TCA Ser	TTA Leu	AGA Arg	AAT Asn	AAA Lys	AAA Lys	1658
GGT Gly	ATT Ile	AGT Ser	GGA Gly	TTC Phe	CTA Leu	AAT Asn	ATT Ile	TTT Phe	GAA Glu	AGT Ser	ATT Ile	TAT Tyr	GAT Asp	AAA Lys	TTT Phe	1706
GGG Gly	GCT Ala	GGT Gly	CAC His	AGT Ser	TCC Ser	ACT Thr	TCA Ser	TTA Leu	AGT Ser	GCT Ala	ATA Ile	CAA Gln	GGA Gly	TAT Tyr	TAT Tyr	1754
GAA Glu	GCC Ala	GAG Glu	TGG Trp	CAA Gln	GTG Val	AAG Lys	AAT Asn	AAA Lys	GAA Glu	AAA Lys	TAT Tyr	GGA Gly	AAT Asn	GGA Gly	GAT Asp	1802
ATA Ile	GAA Glu	ATA Ile	AGT Ser	GAT Asp	AAC Asn	GCA Ala	AAT Asn	GTC Val	ACG Thr	AAT Asn	AAT Asn	GAA Glu	AGG Arg	ATA Ile	TTT Phe	1850
CAA Gln	AAA Lys	GGA Gly	ATA Ile	CAC His	AAT Asn	GAT Asp	AAT Asn	AAT Asn	ATT Ile	AAC Asn	AAT Asn	AAT Asn	ATT Ile	AAT Asn	AAT Asn	1898
AAT Asn	AAT Asn	TAT Tyr	ATC Ile	AAT Asn	CCT Pro	TCA Ser	GAT Asp	GTG Val	GTA Val	GGA Gly	AGA Arg	GAA Glu	AAT Asn	ACG Thr	AAT Asn	1946
GTA Val	CCA Pro	AAT Asn	GTA Val	CGA Arg	AAT Asn	GAT Asp	AAC Asn	CAT His	AAC Asn	GTG Val	GAT Asp	AAA Lys	GTA Val	CAC His	ATT Ile	1994
GCT Ala	ATT Ile	ATA Ile	GGA Gly	GAT Asp	GGT Gly	GGT Gly	TTA Leu	ACA Thr	GGT Gly	GGA Gly	ATG Met	GCA Ala	TTA Leu	GAA Glu	GCG Ala	2042
TTA Leu	AAT Asn	TAT Tyr	ATT Ile	TCA Ser	TTC Phe	TTG Leu	AAT Asn	TCT Ser	AAA Lys	ATT Ile	TTA Leu	ATT Ile	ATT Ile	TAT Tyr	AAT Asn	2090
GAT Asp	AAC Asn	GGA Gly	CAA Gln	GTT Val	TCT Ser	TTA Leu	CCA Pro	ACA Thr	AAT Asn	GCC Ala	GTA Val	AGT Ser	ATA Ile	TCA Ser	GGT Gly	2138
AAT Asn	AGA Arg	CCT Pro	ATA Ile	GGT Gly	TCT Ser	ATA Ile	TCA Ser	GAT Asp	CAT His	TTA Leu	CAT His	TAT Tyr	TTT Phe	GTT Val	TCT Ser	2186

AAT	ATA	GAA	GCA	AAT	GCT	GGT	GAT	AAT	AAA	TTA	TCG	AAA	AAT	GCA	AAA	2234
Asn	Ile	Glu	Ala	Asn	Ala	Gly	Asp	Asn	Lys	Leu	Ser	Lys	Asn	Ala	Lys	
		690					695					700				
GAG	AAT	AAC	ATT	TTT	GAA	AAT	TTG	AAT	TAT	GAT	TAT	ATT	GGT	GTT	GTG	2282
Glu	Asn	Asn	Ile	Phe	Glu	Asn	Leu	Asn	Tyr	Asp	Tyr	Ile	Gly	Val	Val	
	705					710					715					
AAT	GGT	AAT	AAT	ACA	GAA	GAG	CTC	TTT	AAA	GTA	TTA	AAT	AAT	ATA	AAA	2330
Asn	Gly	Asn	Asn	Thr	Glu	Glu	Leu	Phe	Lys	Val	Leu	Asn	Asn	Ile	Lys	
	720				725					730					735	
GAA	AAT	AAA	TTA	AAA	AGA	GCT	ACT	GTT	CTT	CAT	GTA	CGT	ACA	AAA	AAA	2378
Glu	Asn	Lys	Leu	Lys	Arg	Ala	Thr	Val	Leu	His	Val	Arg	Thr	Lys	Lys	
			740						745					750		
TCG	AAT	GAT	TTT	ATA	AAT	TCA	AAG	AGT	CCA	ATA	AGT	ATA	TTG	CAC	TCT	2426
Ser	Asn	Asp	Phe	Ile	Asn	Ser	Lys	Ser	Pro	Ile	Ser	Ile	Leu	His	Ser	
			755					760					765			
ATA	AAG	AAA	AAT	GAG	ATT	TTC	CCT	TTC	GAT	ACC	ACT	ATA	TTA	AAT	GGA	2474
Ile	Lys	Lys	Asn	Glu	Ile	Phe	Pro	Phe	Asp	Thr	Thr	Ile	Leu	Asn	Gly	
		770					775					780				
AAT	ATT	CAT	AAG	GAG	AAC	AAG	ATA	GAA	GAA	GAG	AAA	AAT	GTG	TCT	TCA	2522
Asn	Ile	His	Lys	Glu	Asn	Lys	Ile	Glu	Glu	Glu	Lys	Asn	Val	Ser	Ser	
	785					790					795					
TCT	ACA	AAG	TAT	GAT	GTA	AAT	AAT	AAG	AAT	AAT	AAA	AAT	AAT	GAT	AAT	2570
Ser	Thr	Lys	Tyr	Asp	Val	Asn	Asn	Lys	Asn	Asn	Lys	Asn	Asn	Asp	Asn	
	800				805						810				815	
AGT	GAA	ATT	ATA	AAA	TAT	GAA	GAT	ATG	TTT	TCA	AAA	GAG	ACG	TTC	ACA	2618
Ser	Glu	Ile	Ile	Lys	Tyr	Glu	Asp	Met	Phe	Ser	Lys	Glu	Thr	Phe	Thr	
				820				825						830		
GAT	ATA	TAT	ACA	AAT	GAA	ATG	TTA	AAA	TAT	TTA	AAG	AAA	GAT	AGA	AAT	2666
Asp	Ile	Tyr	Thr	Asn	Glu	Met	Leu	Lys	Tyr	Leu	Lys	Lys	Asp	Arg	Asn	
			835					840					845			
ATA	ATA	TTC	CTA	TCT	CCC	GCT	ATG	TTA	GGA	GGA	TCA	GGA	TTG	GTT	AAA	2714
Ile	Ile	Phe	Leu	Ser	Pro	Ala	Met	Leu	Gly	Gly	Ser	Gly	Leu	Val	Lys	
		850					855					860				
ATT	AGT	GAG	CGT	TAT	CCA	AAT	AAT	GTA	TAT	GAT	GTA	GGT	ATA	GCA	GAA	2762
Ile	Ser	Glu	Arg	Tyr	Pro	Asn	Asn	Val	Tyr	Asp	Val	Gly	Ile	Ala	Glu	
	865					870					875					
CAA	CAT	TCT	GTA	ACT	TTC	GCA	GCA	GCT	ATG	GCA	ATG	AAT	AAG	AAA	TTA	2810
Gln	His	Ser	Val	Thr	Phe	Ala	Ala	Ala	Met	Ala	Met	Asn	Lys	Lys	Leu	
	880				885					890					895	
AAA	ATA	CAA	TTA	TGT	ATA	TAT	TCG	ACC	TTT	TTA	CAA	AGA	GCA	TAT	GAT	2858
Lys	Ile	Gln	Leu	Cys	Ile	Tyr	Ser	Thr	Phe	Leu	Gln	Arg	Ala	Tyr	Asp	
				900				905						910		
CAA	ATT	ATA	CAT	GAT	CTT	AAT	TTA	CAA	AAT	ATA	CCT	TTA	AAG	GTT	ATA	2906
Gln	Ile	Ile	His	Asp	Leu	Asn	Leu	Gln	Asn	Ile	Pro	Leu	Lys	Val	Ile	
			915					920					925			
ATT	GGA	AGA	AGT	GGA	TTA	GTA	GGA	GAG	GAT	GGG	GCA	ACA	CAT	CAA	GGT	2954
Ile	Gly	Arg	Ser	Gly	Leu	Val	Gly	Glu	Asp	Gly	Ala	Thr	His	Gln	Gly	
		930					935					940				
ATA	TAT	GAT	TTA	TCT	TAT	CTT	GGG	ACA	CTT	AAC	AAT	GCA	TAT	ATA	ATA	3002
Ile	Tyr	Asp	Leu	Ser	Tyr	Leu	Gly	Thr	Leu	Asn	Asn	Ala	Tyr	Ile	Ile	
	945					950						955				
TCT	CCA	AGT	AAT	CAA	GTT	GAT	TTG	AAA	AGA	GCT	CTT	AGG	TTT	GCT	TAT	3050
Ser	Pro	Ser	Asn	Gln	Val	Asp	Leu	Lys	Arg	Ala	Leu	Arg	Phe	Ala	Tyr	
	960				965					970					975	
TTA	GAT	AAG	GAC	CAT	TCT	GTG	TAT	ATA	CGT	ATA	CCC	AGA	ATG	AAC	ATA	3098
Leu	Asp	Lys	Asp	His	Ser	Val	Tyr	Ile	Arg	Ile	Pro	Arg	Met	Asn	Ile	
				980					985					990		

TTA AGT GAT AAG TAC ATG AAA GGA TAT TTG AAC ATT CAT ATG AAA AAT	3146
Leu Ser Asp Lys Tyr Met Lys Gly Tyr Leu Asn Ile His Met Lys Asn	
995 1000 1005	
GAG AGC AAA AAT ATC GAT GTA AAC GTG GAT ATA AAC GAT GAT GTA GAT	3194
Glu Ser Lys Asn Ile Asp Val Asn Val Asp Ile Asn Asp Asp Val Asp	
1010 1015 1020	
AAA TAT AGT GAA GAA TAT ATG GAC GAT GAT AAT TTT ATA AAA TCG TTT	3242
Lys Tyr Ser Glu Glu Tyr Met Asp Asp Asp Asn Phe Ile Lys Ser Phe	
1025 1030 1035	
ATT GGA AAA TCT AGA ATT ATT AAA ATG GAT AAT GAA AAT AAT AAT ACA	3290
Ile Gly Lys Ser Arg Ile Ile Lys Met Asp Asn Glu Asn Asn Asn Thr	
1040 1045 1050 1055	
AAT GAA CAT TAT TCA AGC AGA GGA GAT ACA CAG ACA AAA AAA AAA AAA	3338
Asn Glu His Tyr Ser Ser Arg Gly Asp Thr Gln Thr Lys Lys Lys Lys	
1060 1065 1070	
GTT TGT ATC TTT AAC ATG GGT AGT ATG CTT TTT AAT GTA ATT AAT GCT	3386
Val Cys Ile Phe Asn Met Gly Ser Met Leu Phe Asn Val Ile Asn Ala	
1075 1080 1085	
ATA AAA GAA ATT GAA AAA GAA CAA TAT ATT TCA CAT AAT TAT TCT TTT	3434
Ile Lys Glu Ile Glu Lys Glu Gln Tyr Ile Ser His Asn Tyr Ser Phe	
1090 1095 1100	
TCA ATT GTT GAT ATG ATA TTT TTA AAT CCT TTA GAT AAA AAT ATG ATA	3482
Ser Ile Val Asp Met Ile Phe Leu Asn Pro Leu Asp Lys Asn Met Ile	
1105 1110 1115	
GAT CAT GTA ATA AAA CAA AAT AAA CAT CAA TAT TTA ATT ACT TAT GAA	3530
Asp His Val Ile Lys Gln Asn Lys His Gln Tyr Leu Ile Thr Tyr Glu	
1120 1125 1130 1135	
GAT AAT ACT ATA GGT GGT TTT TCT ACA CAT TTC AAT AAT TAT TTA ATA	3578
Asp Asn Thr Ile Gly Gly Phe Ser Thr His Phe Asn Asn Tyr Leu Ile	
1140 1145 1150	
GAA AAT AAT TAT ATT ACA AAA CAT AAC TTA TAT GTT CAT AAT ATT TAT	3626
Glu Asn Asn Tyr Ile Thr Lys His Asn Leu Tyr Val His Asn Ile Tyr	
1155 1160 1165	
TTA TCT AAT GAG CCA ATT GAA CAT GCA TCT TTT AAG GAT CAA CAA GAA	3674
Leu Ser Asn Glu Pro Ile Glu His Ala Ser Phe Lys Asp Gln Gln Glu	
1170 1175 1180	
GTC GTC AAA ATG GAT AAA TGT AGT CTT GTC AAT AGA ATT AAA AAT TAT	3722
Val Val Lys Met Asp Lys Cys Ser Leu Val Asn Arg Ile Lys Asn Tyr	
1185 1190 1195	
CTT AAA AAT AAT CCT ACA TGA TGTAAGATAA ATATATATTT CTAAAATTAT	3773
Leu Lys Asn Asn Pro Thr -	
1200 1205	

TTTTTTTTTA TACTTTAATG TGTACAATAA AATATATATC TAAATATATT TTATTTGTAC3833

GCTTTTTTTTT TTTTTTTTTT AATTGTTATT TTTGTATAT 3872

(4) ANGABEN ZU SEQUENZ ID NR: 4

(i) Sequenzkennzeichen:

(A) LÄNGE: 1205 Aminosäuren

(B) ART: Aminosäure

(ii) Art des Moleküls: PROTEIN

(xi) Sequenzbeschreibung: SEQ ID NO 4:

Met Ile Phe Asn Tyr Val Phe Phe Lys Asn Phe Val Pro Val Val Leu  
 Tyr Ile Leu Leu Ile Ile Tyr Ile Asn Leu Asn Gly Met Asn Asn Lys  
 Asn Gln Ile Lys Thr Glu Lys Ile Tyr Ile Lys Lys Leu Asn Arg Leu  
 Ser Arg Lys Asn Ser Leu Cys Ser Ser Lys Asn Lys Ile Ala Cys Leu

---

Phe Asp Ile Gly Asn Asp Asp Asn Arg Asn Thr Thr Tyr Gly Tyr Asn  
 Val Asn Val Lys Asn Asp Asp Ile Asn Ser Leu Leu Lys Asn Asn Tyr  
 Ser Asn Lys Leu Tyr Met Asp Lys Arg Lys Asn Ile Asn Asn Val Ile  
 Ser Thr Asn Lys Ile Ser Gly Ser Ile Ser Asn Ile Cys Ser Arg Asn  
 Gln Lys Glu Asn Glu Gln Lys Arg Asn Lys Gln Arg Cys Leu Thr Gln  
 Cys His Thr Tyr Asn Met Ser His Glu Gln Asp Lys Leu Ala Asn Asp  
 Asn Asn Arg Asn Asn Lys Lys Asn Phe Asn Leu Leu Phe Ile Asn Tyr  
 Phe Asn Leu Lys Arg Met Lys Asn Ser Leu Leu Asn Lys Asp Asn Phe  
 Phe Tyr Cys Lys Glu Lys Lys Leu Ser Phe Leu His Lys Ala Tyr Lys  
 Lys Lys Asn Cys Thr Phe Gln Asn Tyr Ser Leu Lys Arg Lys Ser Asn  
 Arg Asp Ser His Lys Leu Phe Ser Gly Glu Phe Asp Asp Tyr Thr Asn  
 Asn Asn Ala Leu Tyr Glu Ser Glu Lys Lys Glu Tyr Ile Thr Leu Asn  
 Asn Asn Asn Lys Asn Asn Asn Asn Lys Asn Asn Asp Asn Lys Asn Asn  
 Asp Asn Asn Asp Tyr Asn Asn Asn Asn Ser Cys Asn Asn Leu Gly Glu  
 Arg Ser Asn His Tyr Asp Asn Tyr Gly Gly Asp Asn Asn Asn Pro Cys

Asn Asn Asn Asn Asp Lys Tyr Asp Ile Gly Lys Tyr Phe Lys Gln Ile  
 Asn Thr Phe Ile Asn Ile Asp Glu Tyr Lys Thr Ile Tyr Gly Asp Glu  
 Ile Tyr Lys Glu Ile Tyr Glu Leu Tyr Val Glu Arg Asn Ile Pro Glu  
 Tyr Tyr Glu Arg Lys Tyr Phe Ser Glu Asp Ile Lys Lys Ser Val Leu

Phe Asp Ile Asp Lys Tyr Asn Asp Val Glu Phe Glu Lys Ala Ile Lys  
 Glu Glu Phe Ile Asn Asn Gly Val Tyr Ile Asn Asn Ile Asp Asn Thr  
 Tyr Tyr Lys Lys Glu Asn Ile Leu Ile Met Lys Lys Ile Leu His Tyr  
 Phe Pro Leu Leu Lys Leu Ile Asn Asn Pro Ser Asp Leu Lys Lys Leu  
 Lys Lys Gln Tyr Leu Pro Leu Leu Ala His Glu Leu Lys Ile Phe Leu  
 Phe Phe Ile Val Asn Ile Thr Gly Gly His Phe Ser Ser Val Leu Ser  
 Ser Leu Glu Ile Gln Leu Leu Leu Leu Tyr Ile Phe Asn Gln Pro Tyr  
 Asp Asn Val Ile Tyr Asp Ile Gly His Gln Ala Tyr Val His Lys Ile  
 Leu Thr Gly Arg Lys Leu Leu Phe Leu Ser Leu Arg Asn Lys Lys Gly  
 Ile Ser Gly Phe Leu Asn Ile Phe Glu Ser Ile Tyr Asp Lys Phe Gly  
 Ala Gly His Ser Ser Thr Ser Leu Ser Ala Ile Gln Gly Tyr Tyr Glu  
 Ala Glu Trp Gln Val Lys Asn Lys Glu Lys Tyr Gly Asn Gly Asp Ile  
 Glu Ile Ser Asp Asn Ala Asn Val Thr Asn Asn Glu Arg Ile Phe Gln  
 Lys Gly Ile His Asn Asp Asn Asn Ile Asn Asn Asn Ile Asn Asn Asn  
 Asn Tyr Ile Asn Pro Ser Asp Val Val Gly Arg Glu Asn Thr Asn Val  
 Pro Asn Val Arg Asn Asp Asn His Asn Val Asp Lys Val His Ile Ala

Ile Ile Gly Asp Gly Gly Leu Thr Gly Gly Met Ala Leu Glu Ala Leu  
 Asn Tyr Ile Ser Phe Leu Asn Ser Lys Ile Leu Ile Ile Tyr Asn Asp  
 Asn Gly Gln Val Ser Leu Pro Thr Asn Ala Val Ser Ile Ser Gly Asn  
 Arg Pro Ile Gly Ser Ile Ser Asp His Leu His Tyr Phe Val Ser Asn

---

Ile Glu Ala Asn Ala Gly Asp Asn Lys Leu Ser Lys Asn Ala Lys Glu  
 Asn Asn Ile Phe Glu Asn Leu Asn Tyr Asp Tyr Ile Gly Val Val Asn  
 Gly Asn Asn Thr Glu Glu Leu Phe Lys Val Leu Asn Asn Ile Lys Glu  
 Asn Lys Leu Lys Arg Ala Thr Val Leu His Val Arg Thr Lys Lys Ser  
 Asn Asp Phe Ile Asn Ser Lys Ser Pro Ile Ser Ile Leu His Ser Ile  
 Lys Lys Asn Glu Ile Phe Pro Phe Asp Thr Thr Ile Leu Asn Gly Asn  
 Ile His Lys Glu Asn Lys Ile Glu Glu Glu Lys Asn Val Ser Ser Ser  
 Thr Lys Tyr Asp Val Asn Asn Lys Asn Asn Lys Asn Asn Asp Asn Ser  
 Glu Ile Ile Lys Tyr Glu Asp Met Phe Ser Lys Glu Thr Phe Thr Asp  
 Ile Tyr Thr Asn Glu Met Leu Lys Tyr Leu Lys Lys Asp Arg Asn Ile  
 Ile Phe Leu Ser Pro Ala Met Leu Gly Gly Ser Gly Leu Val Lys Ile  
 Ser Glu Arg Tyr Pro Asn Asn Val Tyr Asp Val Gly Ile Ala Glu Gln  
 His Ser Val Thr Phe Ala Ala Ala Met Ala Met Asn Lys Lys Leu Lys  
 Ile Gln Leu Cys Ile Tyr Ser Thr Phe Leu Gln Arg Ala Tyr Asp Gln  
 Ile Ile His Asp Leu Asn Leu Gln Asn Ile Pro Leu Lys Val Ile Ile

Gly Arg Ser Gly Leu Val Gly Glu Asp Gly Ala Thr His Gln Gly Ile  
Tyr Asp Leu Ser Tyr Leu Gly Thr Leu Asn Asn Ala Tyr Ile Ile Ser  
Pro Ser Asn Gln Val Asp Leu Lys Arg Ala Leu Arg Phe Ala Tyr Leu  
Asp Lys Asp His Ser Val Tyr Ile Arg Ile Pro Arg Met Asn Ile Leu

---

Ser Asp Lys Tyr Met Lys Gly Tyr Leu Asn Ile His Met Lys Asn Glu  
Ser Lys Asn Ile Asp Val Asn Val Asp Ile Asn Asp Asp Val Asp Lys  
Tyr Ser Glu Glu Tyr Met Asp Asp Asp Asn Phe Ile Lys Ser Phe Ile  
Gly Lys Ser Arg Ile Ile Lys Met Asp Asn Glu Asn Asn Asn Thr Asn  
Glu His Tyr Ser Ser Arg Gly Asp Thr Gln Thr Lys Lys Lys Lys Val  
Cys Ile Phe Asn Met Gly Ser Met Leu Phe Asn Val Ile Asn Ala Ile  
Lys Glu Ile Glu Lys Glu Gln Tyr Ile Ser His Asn Tyr Ser Phe Ser  
Ile Val Asp Met Ile Phe Leu Asn Pro Leu Asp Lys Asn Met Ile Asp  
His Val Ile Lys Gln Asn Lys His Gln Tyr Leu Ile Thr Tyr Glu Asp  
Asn Thr Ile Gly Gly Phe Ser Thr His Phe Asn Asn Tyr Leu Ile Glu  
Asn Asn Tyr Ile Thr Lys His Asn Leu Tyr Val His Asn Ile Tyr Leu  
Ser Asn Glu Pro Ile Glu His Ala Ser Phe Lys Asp Gln Gln Glu Val  
Val Lys Met Asp Lys Cys Ser Leu Val Asn Arg Ile Lys Asn Tyr Leu  
Lys Asn Asn Pro Thr



Plasmodium falciparum gcpE-Gen

- (i) SEQUENZKENNZEICHEN:
- (D) LÄNGE: 2109 BASENPAARE
- (E) ART: genomische Sequenz
- (ii) ART DES MOLEKÜLS: DNA
- (iii) URSPRÜNGLICHE HERKUNFT:
- (B) ORGANISMUS: Plasmodium falciparum

FUNKTION: essentielles Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO:5

cagcctataaatattattattattattatttttttttttttttttttttttcataatgcctgaat  
 Q P I N I I I Y Y Y F F F F F S - C L N  
 aaccacaaaatgagttatataaaaagactgattcctttttatgttactgttttattctcat  
 N H K M S Y I K R L I L F M L L F Y S H  
 gtaaaaattaaaaaattattttattaaaatttctaattgtaaacatattttttgcagaagca  
 V K I K K L F I K I S N V N I F F A E A  
 aagaaaaatggaaaaaaggaattcctttcctttttttactaaatataaaaaaaatagccaa  
 K K N G K K E F F L F L L N I K K N S Q  
 cagaaaaaaacttatcatattaccaaaaggaataccataaataaaaagtgttttttatat  
 Q K K T Y H I T K R N T I N K S D F L Y  
 tctttactaaatgaagaagggaattccttcaaaaaaggaatataaaaattttaaaagatgaa  
 S L L N E E G N S S K K E Y K N L K D E  
 gaaaaatataatatcatacaaaatataaaaaaatattgtgaatgtactaaaaaatataaa  
 E K Y N I I Q N I K K Y C E C T K K Y K  
 aggctcccaacacgagaagtagttatttggaatgttaaaattggaggaaataataaaaata  
 R L P T R E V V I G N V K I G G N N K I  
 gctattcaaactatggctagctgtgatacaagaaatgtagaagaatgtgtatatcaaatt  
 A I Q T M A S C D T R N V E E C V Y Q I  
 agaaaatgtaaagatttggtgctgacattgtaagggttgactgttcaaggagttcaagaa  
 R K C K D L G A D I V R L T V Q G V Q E  
 gcacaagctagttatcatattaaagaaaaattattatctgaaaatgtaaatatcccat  
 A Q A S Y H I K E K L L S E N V N I P L  
 gtaacagatattcatttttaattcctaaaatagcttttaattggcagctgatgtgtttgaaaaa  
 V T D I H F N P K I A L M A A D V F E K  
 attcgagtgaatccaggaaattatgttgatggaagaaaaaatggatagataaagtttat  
 I R V N P G N Y V D G R K K W I D K V Y  
 aaaactaaagaagaattttgatgaagggaattattttataaaagaaaaattttgtaccatta  
 K T K E E F D E G K L F I K E K F V P L  
 attgaaaaatgtaaagattaaatagagcaataagaattggtacaaatcatggattcctt  
 I E K C K R L N R A I R I G T N H G F L  
 tcatctcgagtattatcatattatggagatacaccattagcattagtagaaagtgtatg  
 S S R V L S Y Y G D T P L A L V E S A M  
 agattttctgattttatgtaatgaaaacaatttttaacaatcttgttttttctatgaaagct  
 R F S D L C N E N N F N N L V F S M K A  
 tctaattgcttatgttatgatacaatcttatagattattagtagtatctaacaatatgaaaga  
 S N A Y V M I Q S Y R L L V S K Q Y E R  
 aatatgatgttccctatacatatttaggagttacagaagcaggatttggtgataatggaaga  
 N M M F P I H L G V T E A G F G D N G R

```

ataaaatcttatttaggtataggatctttatttatatgatgggtataggagataccattcgt
I K S Y L G I G S L L Y D G I G D T I R
atatccttaacagaagatccttggaagagtttaactccttgtaaaaaattagttgaaaat
I S L T E D P W E E L T P C K K L V E N
ttaaagaaaaagaatatattttataatgaaaatttttaagaagataatgaattaaaaataat
L K K R I F Y N E N F K E D N E L K N N
gaaatggataccaaaaatctattaaattttgaagaaaattatcgaaattttaataatata
E M D T K N L L N F E E N Y R N F N N I
aaaaaaagaaatgtagaaaaaaataataatgtattacatgaagagtgactataggtaat
K K R N V E K N N N V L H E E C T I G N
gtagtaaccataaaagagtttagaagattctctgcaaatttttaagatttaaattagaa
V V T I K E L E D S L Q I F K D L N L E
gtagattcaaattggaatttgaaaaaggagccaaaacaactgatatgggtattataaat
V D S N G N L K K G A K T T D M V I I N
gattttcataatataacaaatttaggaaaaaaaactgtggataaattaatgcaagtggga
D F H N I T N L G K K T V D K L M Q V G
attaatatagtagttcaatatgaaccacataatatagaatttatagaaaaaatggaacca
I N I V V Q Y E P H N I E F I E K M E P
aataatgataataataataataataataataataatattattttatgtggatataaaa
N N D N N N N N N N N N I L F Y V D I K
aatattatgaacagttcagaaaaaaatattaaattaagtaattctaaaggatatggatta
N I M N S S E K N I K L S N S K G Y G L
attttaaacggaaaagaagatatatacaaacataaaaaaaataaaagaattaaatcgtcgt
I L N G K E D I Q T I K K I K E L N R R
cctttattcattctattaaaatcagataacatatatgaacatgtattaataaccagaaga
P L F I L L K S D N I Y E H V L I T R R
attaatgaacttttacaatccttaaatataaatataaccttatatacattatgttgatatt
I N E L L Q S L N I N I P Y I H Y V D I
aattcaaataattatgatgatatatattagttaattcaacattatatgcaggaagttgtttg
N S N N Y D D I L V N S T L Y A G S C L
atggatttaattgggggatgggtcttattgttaacgtaactaatgatgttcttacaataaa
M D L M G D G L I V N V T N D V L T N K
aaagggtag
K G -

```

- (6) SEQ ID NO: 6
- (i) Sequenzkennzeichen:
  - (A) Länge: 679 Aminosäuren
  - (B) Art: Aminosäure
- (ii) Art des Moleküls: Protein
- (iv) URSPRÜNGLICHE HERKUNFT:
  - (A) ORGANISMUS: Plasmodium falciparum

```

1  MSYIKRLILF MLLFYSHVKI KKLFIKISNV NIFFAEAKKN GKKEFFLFLL
51  NIKKNSQQKK TYHITKRNTI NKSDFLYSLN NEEGNSSKKE YKNLKDDEKY
101 NIIQNIKKYC ECTKKYKRLP TREVVIGNVK IGGNNKIAIQ TMASCDTRNV
151 EECVYQIRKC KDLGADIVRL TVQGVQEAQA SYHIKEKLLS ENVNIPLVTD

```

201	IHFNPKIALM	AADVFEKIRV	NPGNYVDGRK	KWIDKVYKTK	EEFDEGKLFI
251	KEKFVPLIEK	CKRLNRAIRI	GTNHGFLSSR	VLSYYGDTPL	ALVESAMRFS
301	DLCNENNFNN	LVFSMKASNA	YVMIQSYRLL	VSKQYERNMM	FPIHLGVTEA
351	GFGDNGRIKS	YLGIGSLLYD	GIGDTIRISL	TEDPWEELTP	CKKLVENLKK
401	RIFYNENFKE	DNELKNNEMD	TKNLLNFEEN	YRNFNNIKKR	NVEKNNNVLH
451	EECTIGNVVT	IKELEDSLQI	FKDLNLEVDS	NGNLKKGAKT	TDMVIINDFH
501	NITNLGKKTV	DKLMQVGINI	VVQYEPHNIE	FIEKMEPNND	NNNNNNNNNI
551	LFYVDIKNIM	NSSEKNIKLS	NSKGYGLILN	GKEDIQTIKK	IKELNRRPLF
601	ILLKSDNIYE	HVLITRRINE	LLQSLNINIP	YIHYVDINSN	NYDDILVNST
651	LYAGSCLMDL	MGDGLIVNVT	NDVLTNKKG-		

---

Patentansprüche

1. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 2 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:2, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.
2. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 4 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:4, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.
3. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO:6, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.
4. DNA-Sequenz gemäß einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß sie außerdem funktionelle Regulations-signale, insbesondere Promotoren, Operatoren, Enhancer, ribosomale Bindungsstellen, aufweist.
5. DNA-Sequenz mit folgenden Teilsequenzen
  - i) Promotor, der in Viren, Eukaryonten und Prokaryonten aktiv ist und die Bildung einer RNA im vorgesehenen Zielgewebe oder den Zielzellen sicherstellt,
  - ii) DNA-Sequenzen gemäß einem der Ansprüche 1 bis 3,

iii) 3'-nichttranslatierte Sequenz, die in Viren, Eukaryonten und Prokaryonten zur Addition von Poly-A Resten an das 3'-Ende der RNA führt.

6. Verfahren zur Herstellung von transgenen Viren, Eukaryonten und Prokaryonten zur Veränderung des Isoprenoid-Gehaltes, dadurch gekennzeichnet, daß eine DNA-Sequenz gemäß Anspruch 4 oder 5 in das Genom von Viren, eukaryontischen und prokaryontischen Zellen mit oder ohne Verwendung eines Plasmids transferiert und eingebaut wird.
7. Expressionsvektor, enthaltend eine oder mehrere DNA-Sequenzen gemäß Anspruch 1 bis 3.
8. Protein, welches am 1-Deoxy-D-Xylulose-5-Phosphat-Stoffwechselweges beteiligt ist und a) codiert wird von der in DNA-Sequenz SEQ ID NO: 1,3,5 oder b) codiert wird von DNA-Sequenzen, die mit den DNA-Sequenzen SEQ ID NO: 1,3,5 oder Fragmenten dieser DNA-Sequenzen im DNA-Bereich, der für das reife Protein codiert, hybridisieren.
9. Protein nach den Anspruch 8, erhältlich aus den Kulturüberständen von Parasiten oder aus den aufgeschlossenen Parasiten und Aufreinigung über chromatographische und elektrophoretische Techniken.
10. Protein nach einem der Ansprüche 8 und 9, dadurch gekennzeichnet, daß es a) das Produkt einer viralen, prokaryontischen oder eukaryontischen Expression einer exogenen DNA ist, b) codiert wird von den Sequenzen SEQ ID NO: 1, 3 oder 5 oder codiert wird von DNA-Sequenzen, die mit den in den DNA-Sequenzen SEQ ID NO: 1, 3 oder 5 oder Fragmenten dieser DNA-Sequenzen im DNA-Bereich, der für das reife Protein codiert, hybridisieren, oder c) codiert wird von DNA-Sequenzen, die ohne Degeneration des genetischen Codes mit den in b) definierten Sequenzen hybridisieren würden und

für ein Polypeptid mit entsprechender Aminosäure-Sequenz kodieren.

11. Protein gemäß einem der vorangehenden Ansprüchen, welches aus den Aminosäuren von Sequenz SEQ ID NO: 2, 4 und 6 besteht.
- 
12. Verfahren zur Bestimmung der enzymatische Aktivität des gcpE-Proteins, dadurch gekennzeichnet, daß Phosphorylierung eines Zuckers oder eines Phosphorzuckers oder einer Vorstufe der Isoprenoidbiosynthese, insbesondere die Phosphorylierung von 2-C-Methyl-D-erythritol, 2-C-Methyl-D-erythritol-phosphat, insbesondere 2-C-Methyl-D-erythritol-4-phosphat, 2-C-Methyl-D-erythrose, 2-C-Methyl-D-erythrose-phosphat, insbesondere 2-C-Methyl-D-erythrose-4-phosphat, und der Phosphat- und Alkoholvorstufen, detektiert wird.
  13. Verfahren nach Anspruch 12, dadurch gekennzeichnet, daß die Phosphorylierung der folgenden Phosphate oder Alkohole detektiert wird:
 
$$\begin{aligned} &\text{CH}_2(\text{OH})-\text{C}(\text{CH}_3)=\text{C}(\text{OH})-\text{CH}_2-\text{O}-\text{PO}(\text{OH})_2, \\ &\text{CH}_2(\text{OH})-\text{C}(\text{CH}_3)=\text{C}(\text{OH})-\text{CH}_2-\text{OH}, \\ &\text{CH}_2(\text{OH})-\text{CH}(\text{CH}_3)-\text{CO}-\text{CH}_2-\text{O}-\text{PO}(\text{OH})_2, \text{CH}_2(\text{OH})-\text{CH}(\text{CH}_3)-\text{CO}-\text{CH}_2-\text{OH} \\ &\text{CH}_2=\text{C}(\text{CH}_3)-\text{CO}-\text{CH}_2-\text{O}-\text{PO}(\text{OH})_2, \text{CH}_2=\text{C}(\text{CH}_3)-\text{CO}-\text{CH}_2-\text{OH}, \\ &\text{CH}_2=\text{C}(\text{CH}_3)-\text{CH}(\text{OH})-\text{CH}_2-\text{O}-\text{PO}(\text{OH})_2, \text{CH}_2=\text{C}(\text{CH}_3)-\text{CH}(\text{OH})-\text{CH}_2-\text{OH}, \\ &\text{CH}_2(\text{OH})-\text{C}(=\text{CH}_2)-\text{C}(\text{OH})-\text{CH}_2-\text{O}-\text{PO}(\text{OH})_2, \\ &\text{CH}_2(\text{OH})-\text{C}(=\text{CH}_2)-\text{C}(\text{OH})-\text{CH}_2-\text{OH} \\ &\text{CHO}-\text{CH}(\text{CH}_3)-\text{CH}(\text{OH})-\text{CH}_2-\text{O}-\text{PO}(\text{OH})_2, \text{CHO}-\text{CH}(\text{CH}_3)-\text{CH}(\text{OH})-\text{CH}_2-\text{OH}, \\ &\text{CH}_2(\text{OH})-\text{C}(\text{OH})(\text{CH}_3)-\text{CH}=\text{CH}-\text{O}-\text{PO}(\text{OH})_2, \\ &\text{CH}_2(\text{OH})-\text{C}(\text{OH})(\text{CH}_3)-\text{CH}=\text{CH}-\text{OH} \\ &\text{CH}(\text{OH})=\text{C}(\text{CH}_3)-\text{CH}(\text{OH})-\text{CH}_2-\text{O}-\text{PO}(\text{OH})_2, \\ &\text{CH}(\text{OH})=\text{C}(\text{CH}_3)-\text{CH}(\text{OH})-\text{CH}_2-\text{OH} \end{aligned}$$
  14. Verfahren zum Screening einer Verbindung, wobei das Verfahren umfaßt:
    - a) Bereitstellen einer Wirtszelle, die einen rekombinanten Expressionsvektor enthält, wobei der Vektor zumindest einen Teil der Oligonukleotidsequenz gemäß SEQ ID NO:1,

SEQ ID NO: 3 oder SEQ ID NO: 5 oder Varianten oder Analoga dieser aufweist, und außerdem eine Verbindung, von der vermutet wird, daß sie eine antimykotische, antibiotische, antiparasitäre oder antivirale Wirkung bei Mensch und Tier hat,

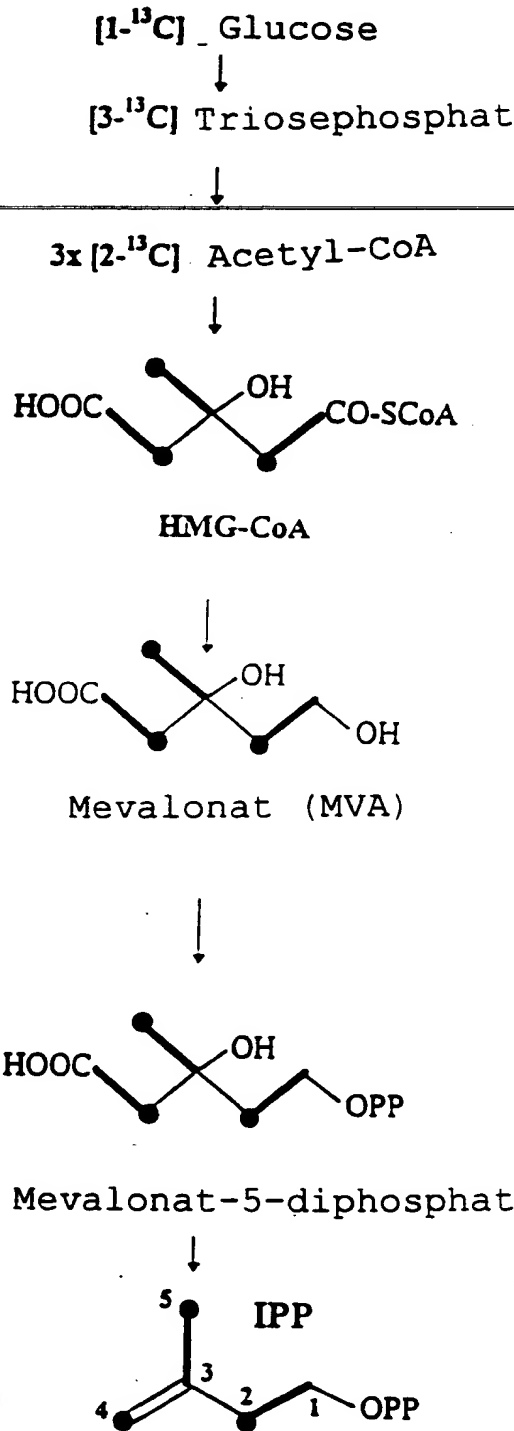
- b) In-Kontakt-Bringen der Wirtszelle mit der Verbindung und
  - c) Bestimmung der antimikrobiellen, antimykotischen, antibiotischen, antiparasitären oder antiviralen Wirksamkeit der Verbindung.
- 

15. Verfahren zum Screening einer Verbindung, wobei das Verfahren umfaßt:

- a) Bereitstellen einer Wirtszelle, die einen rekombinanten Expressionsvektor enthält, wobei der Vektor zumindest einen Teil der Oligonukleotidsequenz gemäß SEQ ID NO:1, SEQ ID NO: 3 oder SEQ ID NO: 5 oder Varianten oder Analoga dieser aufweist, und außerdem eine Verbindung, von der vermutet wird, daß sie eine antimikrobielle, antivirale, antiparasitäre, bakterizide, fungizide oder herbizide Wirkung bei Pflanzen hat,
- b) In-Kontakt-Bringen der Wirtszelle mit der Verbindung und
- c) Bestimmung der antimikrobiellen, antiviralen, antiparasitären, bakteriziden, fungiziden oder herbiziden Wirksamkeit der Verbindung.

11.05.1999

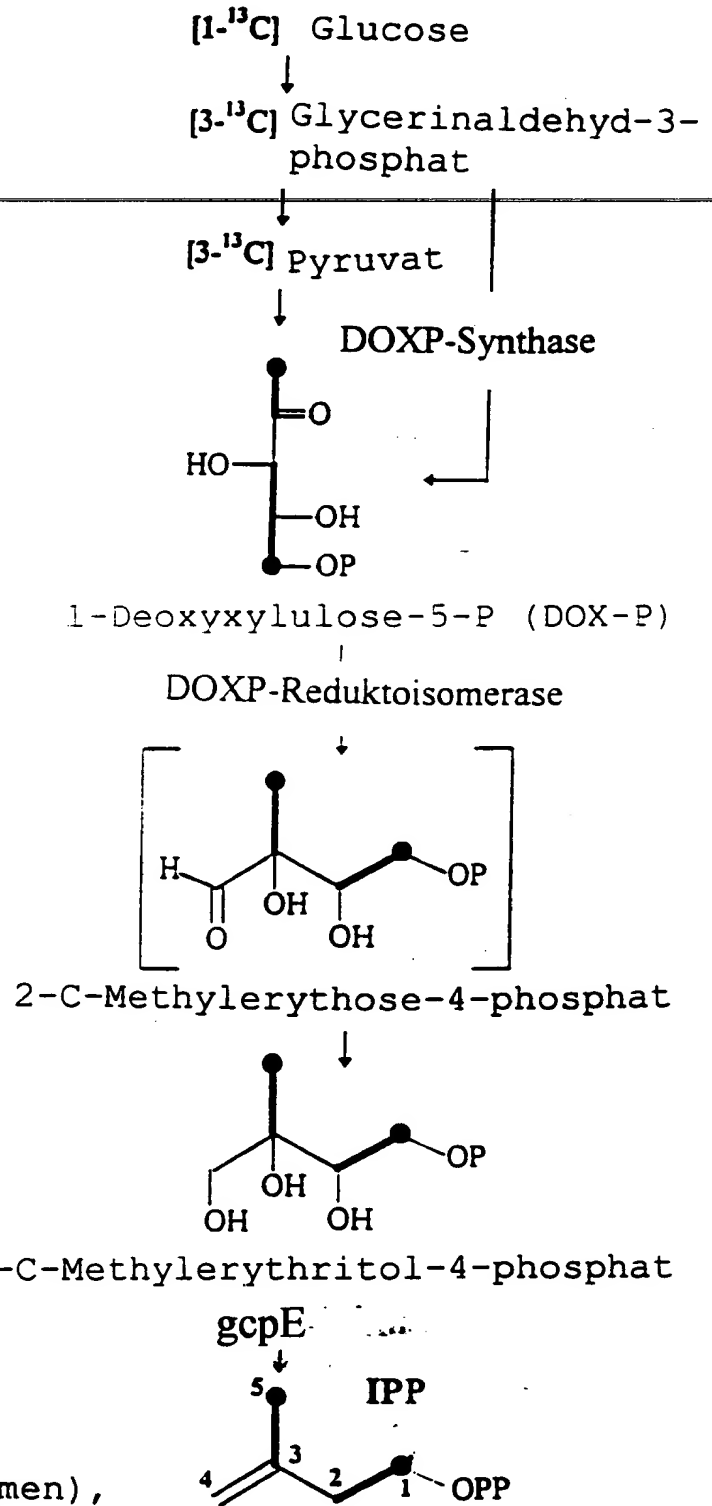
# Klassischer Acetat/ Mevalonat-Pathway



höhere Pflanzen (Cytoplasmen),  
Tiere, Pilze, Eubakterien

Fig. 1

# Alternativer DOX-P Pathway



höhere Pflanzen (Plastide),  
Grünalgen, viele Eubakterien